

Appendix A: Example Calculations of Statistical Methods

1.0 Example 1

Assume 12 baseline iron loading observations are collected by sampling once per month for a year. Likewise, 12 iron loading monitoring observations are obtained by sampling once per month for a period of one year. In order to determine whether baseline pollution loading has been exceeded, both Procedures A and B were used. For all calculations in Example 1, assume the following iron loading observations (lbs/day).

Baseline	1.33	0.53	0.92	0.82	0.88	0.79	0.87	0.73	0.83	0.89	1.10	0.86
Monitoring	0.94	0.74	0.87	1.03	0.91	1.00	0.80	1.19	1.16	1.15	1.12	0.91

1.1 Procedure A (Figure 3.2a)

1.1.1 Single Observation Trigger:

- 1) Twelve baseline observations were collected, therefore $n = 12$.
- 2) The baseline observations were placed in sequential order from smallest to largest.
[0.53, 0.73, 0.79, 0.82, 0.83, 0.86, 0.87, 0.88, 0.89, 0.92, 1.10, 1.33]
- 3) The number of observations, n , is less than 16, therefore the Single Observation Trigger (L) equals $x_{(12)}$, (the maximum) = 1.33.
- 4) All monitoring observations are less than 1.33, therefore the Single Observation Trigger (L) (1.33) was not exceeded.

1.1.2 Subtle Trigger

- 1) Twelve is an even number, therefore the median of the baseline observations is:

$$M = 0.5 * (x_{(6)} + x_{(7)}).$$

$$M = 0.5 * (0.86 + 0.87) = 0.865$$

In order to determine M_1 , calculate the median of the subset ranging from $x_{(7)}$ to $x_{(12)}$.
Because $12 - 6 = 6$ is even, $M_1 = 0.5 * (x_{(9)} + x_{(10)})$
 $M_1 = 0.5 * (0.89 + 0.92) = 0.905$

2) In order to determine M_{-1} , calculate the median of the subset ranging from $x_{(1)}$ to $x_{(6)}$.
Because 6 is even, $M_{-1} = 0.5 * (x_{(3)} + x_{(4)})$
 $M_{-1} = 0.5 * (0.79 + 0.82) = 0.805$

3) To calculate R, subtract M_{-1} from M_1 .
 $R = 0.905 - 0.805 = 0.1$

4) The calculated value for R, is then substituted into the equation for T.

$$T = 0.865 + \frac{1.58 * [(1.25 * 0.1)]}{(1.35 * \sqrt{12})} = 0.907$$

5) The following monitoring observations are ordered from smallest to largest.
[0.74, 0.80, 0.87, 0.91, 0.91, 0.94, 1.00, 1.03, 1.12, 1.15, 1.16, 1.19]

6) There are 12 monitoring observations, therefore $m = 12$.
The number of observations is even, therefore $M' = 0.5 * (x_{(6)} + x_{(7)})$
 $M' = 0.5 * (0.94 + 1.00) = 0.97$
This holds true for M_1' and M_{-1}' as well.
 $M_1' = 0.5 * (x_{(9)} + x_{(10)}) = 0.5 * (1.12 + 1.15) = 1.135$
 $M_{-1}' = 0.5 * (x_{(3)} + x_{(4)}) = 0.5 * (0.87 + 0.91) = 0.89$

7) To calculate R, subtract M_{-1}' from M_1' .
 $R' = 1.135 - 0.89 = 0.245$.

8) The calculated value for R' is then substituted in the equation for T'.

$$T' = 0.925 - \frac{1.58 * [(1.25 * 0.245)]}{(1.35 * \sqrt{12})} = 0.867$$

9) T' (0.867) is less than T (0.907), therefore the median baseline pollution loading was not exceeded.

1.2 Procedure B (Figure 3.2b)

1.2.1 Calculation of Single Observation Limit

- 1) Again, the number of baseline observations, $n = 12$.
- 2) The following log-transformed (using natural logs) baseline observations are sequentially ordered from smallest to largest. [-0.64, -0.31, -0.24, -0.20, -0.19, -0.15, -0.14, -0.13, -0.12, -0.08, 0.10, 0.29]
- 3) The mean of the 12 log-transformed observations, $E_y = -0.15$.
- 4) An appropriate estimate of the first-order autocorrelation (ρ_1) of the log-transformed data is 0.5. Given the number of observations and the auto-correlation estimate, the following equation is used to calculate A.

$$A = \frac{1}{[1 - (\frac{2}{12}) * 0.5]} = 1.09$$

- 5) The factor A is then used to calculate S_y^2 .

$$S_y^2 = 1.09 * \sum \frac{[y_i - (-0.15)]^2}{n-1} = 0.0535$$

- 6) To find E_x the values for E_y and S_y^2 into the following equation:

$$E_x = \exp [(E_y) + (0.5 * S_y^2)]$$

$$E_x = \exp [(-0.15) + (0.5 * 0.0535)] = \exp (-0.12325) = 0.884$$

- 7) The Single Observation Limit (L_{so}) is defined as the following:

$$L_{so} = \exp [(E_y) + (Z_{99} * \sqrt{S_y^2})]$$

$$L_{so} = \exp [(-0.15) + (2.3263 * \sqrt{0.0535})]$$

$$L_{so} = \exp [0.388] = 1.47$$

- 8) Monitoring observations are below 1.47, therefore the L_{so} was not exceeded.

1.2.2 Calculation of Single Observation Warning Level

1) The Single Observation Warning Level (WL_{so}) is determined by using the following equation:

$$WL_{so} = \exp [(-0.15) + (1.6449 * \sqrt{0.0535})] = \exp[0.230] = 1.26$$

2) All of the monitoring observations are below 1.26, therefore the Single Observation Warning Level was not exceeded.

1.2.3 Calculation of Cusum test

1) The number of monitoring observations, $n = 12$.

2) The log-transformed (using natural logs) monitoring observations are listed and labeled sequentially, in order of collection.

	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}	Y_{11}	Y_{12}
Obs.	-0.06	-0.30	-0.14	0.03	-0.09	0.00	-0.22	0.17	0.15	0.14	0.11	-0.09

3) Using the values for E_y and S_y^2 from the L_{so} calculations, the value for K can be determined using the following equation:

$$K = [(E_y) + 0.25(\sqrt{S_y^2})]$$

$$K = (-0.15) + 0.25 * (0.229) = -0.092$$

4) The values for $C_{(i)}$, can be determined using the following equation:

$$C_t = C_{t-1} + (Y_n - K) \text{ for example}$$

$$C_1 = 0 + (-0.06 - (-0.093)) = 0.033$$

The values for C_t are given in the following table for each collection time t . Negative C_t values are replaced with 0, as shown in parentheses.

t	C_t	t	C_t
1	0.033	7	0.092
2	-0.174 (0)	8	0.355
3	-0.047 (0)	9	0.598
4	0.123	10	0.831
5	0.126	11	1.034
6	0.219	12	1.037

- 5) The baseline pollution Cusum Single Observation Limit, H , can be determined using the following equation:

$$H = 8.0 * (\sqrt{S_y^2})$$

$$H = 8.0 * 0.229 = 1.850$$

- 6) All values for C are below 1.850, therefore the baseline pollution Cusum Single Observation Limit was not exceeded.

1.2.4 Cusum Warning Level

- 1) The number of monitoring observations, $n = 12$.

- 2) The following equation can be used to determine K_w :

$$K_w = [(E_y) + 0.5(\sqrt{S_y^2})]$$

$$K_w = [(-0.15) + 0.5(0.229)] = -0.0393$$

- 3) The values for W_t , can be determined using the following equation:

$$W_t = W_{t-1} + (Y_n - (K_w)) \text{ for example}$$

$$W_1 = 0 + (-0.06 - (-0.0355)) = -0.0245 = 0.00$$

The limit values, W_t , are given in the following table for each collection time t . Negative W_t values were replaced with 0, as shown in parentheses.

t	W_t	t	W_t
1	-0.0257 (0)	7	-0.1428 (0)
2	-0.2657 (0)	8	0.2043
3	-0.1057 (0)	9	0.3886
4	0.0643	10	0.5629
5	0.0086	11	0.7072
6	0.0429	12	0.6515

5) The baseline pollution Cusum Warning Level, H_w , can be determined using the following equation:

$$H_w = 3.5 * (\sqrt{S_y^2})$$

$$H_w = 3.5 * (0.229) = 0.8096$$

6) All values for W_t are below 0.8096, therefore the baseline pollution Cusum Warning Level was not reached or exceeded.

1.3 Annual Comparisons

1.3.1 Wilcoxon-Mann-Whitney Test

Instructions for the Wilcoxon-Mann-Whitney test are given in Conover (1980), cited in Figure 3.2b.

- 1) When using both baseline and monitoring data, $n = 12$ and $m = 12$
- 2) The baseline and monitoring observations are listed with their corresponding rankings.

Baseline Observations	0.53	0.73	0.79	0.82	0.83	0.86	0.87	0.88	0.89	0.92	1.10	1.33
Baseline Rankings	1	2	4	6	7	8	9.5	11	12	15	19	24

Monitoring Observations	0.74	0.80	0.87	0.91	0.91	0.94	1.00	1.03	1.12	1.15	1.16	1.19
Monitoring Rankings	3	5	9.5	13.5	13.5	16	17	18	20	21	22	23

Due to the fact that values of 0.87 and 0.91 were each obtained for more than one observation. The average rankings are obtained for these values. For 0.87, the average of 9 and 10 is 9.5. For 0.91, the average of 13 and 14 is 13.5.

- 3) The sum of the twelve baseline ranks, $S_n = 118.5$.
- 4) In order to find the appropriate critical value (C), match the column with the correct n (number of baseline observations) to the row with the correct m (number of monitoring observations). As found in the table, the critical value C for 12 baseline and 12 monitoring observations is 121.

*Critical Values (C) of the Wilcoxon-Mann-Whitney Test
(for a one-sided test at the 95 percent level)*

n	10	11	12	13	14	15	16	17	18	19	20
m											
10	83	98	113	129	147	165	185	205	227	249	273
11	87	101	117	134	152	171	191	211	233	256	280
12	90	105	121	139	157	176	197	218	240	263	288
13	93	109	126	143	162	182	202	224	247	271	295
14	97	113	130	148	167	187	208	231	254	278	303
15	100	117	134	153	172	193	214	237	260	285	311
16	104	121	139	157	177	198	220	243	267	292	318
17	107	124	143	162	183	204	226	250	274	300	326
18	111	128	147	167	188	209	232	256	281	307	334
19	114	132	151	172	193	215	238	263	288	314	341
20	118	136	156	176	198	221	244	269	295	321	349

- 5) S_n (118.5) is less than C (121). Therefore, according to the Wilcoxon-Mann-Whitney test, the monitoring observations did exceed the baseline pollution loading.

2.0 Example 2

Assume 18 baseline iron loading determination observations are collected by sampling twice per month for nine months. Likewise, 18 iron loading monitoring observations are obtained by sampling twice per month for a period of nine months. In order to determine whether baseline pollution loading has been exceeded, examples of Procedures A and B are presented below. For all calculations in Example 2, assume the following iron loading observations (in lbs/day):

Observation	Baseline	Monitoring
1	0.030	0.530
2	0.005	0.040
3	1.915	1.040
4	0.673	0.033
5	0.064	0.030
6	0.063	0.230
7	0.607	0.710
8	0.553	0.240
9	0.286	0.390
10	0.106	0.830
11	0.406	3.050
12	1.447	0.580
13	0.900	1.180
14	0.040	0.510
15	2.770	0.046
16	1.803	0.690
17	0.160	0.630
18	0.045	0.370

2.1 Procedure A

2.1.1 Single Observation Trigger

- 1) The number of baseline observations collected, $n = 18$.
- 2) The baseline observations are ordered sequentially from smallest to largest.
[0.005, 0.030, 0.040, 0.045, 0.063, 0.064, 0.106, 0.160, 0.286, 0.406, 0.553, 0.607, 0.673, 0.900, 1.447, 1.803, 1.915, 2.770]
- 3) The number of observations is greater than 16, therefore M , M_1 , M_2 and M_3 must be calculated. The number of observations is even, which means the median of the baseline observations must be calculated using the following equation:

$$M = 0.5 * (x_{(9)} + x_{(10)}).$$

$$M = 0.5 * (0.286 + 0.406) = 0.346$$
- 4) To determine M_1 , calculate the median of the subset ranging from $x_{(10)}$ to $x_{(18)}$.
 $18 - 9 = 9$ is odd, therefore $M_1 = x_{(14)} = 0.900$.
- 5) To determine M_2 , calculate the median of the subset ranging from $x_{(14)}$ to $x_{(18)}$.
 $18 - 13 = 5$ is odd, therefore $M_2 = x_{(16)} = 1.803$.
- 6) To determine M_3 , calculate the median of the subset ranging from $x_{(16)}$ to $x_{(18)}$.
 $18 - 15 = 3$, which is odd, therefore $M_3 = x_{(17)} = 1.915$.
- 7) To determine L , calculate the median of the subset ranging from $x_{(17)}$ to $x_{(18)}$.
 $18 - 16 = 2$, which is even, therefore $L = 0.5 * (x_{(17)} + x_{(18)}) = 0.5 * (1.915 + 2.770) = 2.343$.
- 8) One monitoring observation, 3.050, is above L , (2.343), therefore the Single Observation Trigger was exceeded.

2.1.2 Subtle Trigger:

- 1) As determined in section 2.1.1, $M = 0.346$, and $M_1 = 0.900$.
- 2) To find M_{-1} , calculate the median of the subset ranging from $x_{(1)}$ to $x_{(9)}$.
 9 is odd, therefore $M_{-1} = x_{(5)} = 0.063$.
- 3) The value for R is found by subtracting M_{-1} from M
 $R = 0.900 - 0.063 = 0.837$

4) To find T, the value for R is inserted in the following equation:

$$T = 0.346 + \frac{1.58 * [(1.25 * 0.837)]}{(1.35 * \sqrt{18})} = 0.635$$

5) The monitoring observations are placed in order from lowest to highest.

[0.030, 0.033, 0.040, 0.046, 0.230, 0.240, 0.370, 0.390, 0.510, 0.530, 0.580, 0.630, 0.690, 0.710, 0.830, 1.040, 1.180, 3.050]

6) The number of monitoring observations, $m = 18$.

18 is even, making $M' = 0.5 * (x_{(9)} + x_{(10)})$

$$M' = 0.5 * (0.510 + 0.530) = 0.520$$

7) To determine M_1' , calculate the median of subset $x_{(10)}$ to $x_{(18)}$.

Because $18 - 9 = 9$ is odd, $M_1' = (x_{(14)}) = 0.710$

8) To determine M_{-1}' , calculate the median of subset $x_{(1)}$ to $x_{(9)}$.

Because 9 is odd, $M_{-1}' = (x_{(5)}) = 0.230$

9) The value for R' is found by subtracting M_{-1}' from M_1' .

$$R' = 0.710 - 0.230 = 0.48$$

10) To find T' , the value for R' is inserted into the following equation:

$$T' = 0.520 - \frac{1.58 * [(1.25 * 0.48)]}{(1.35 * \sqrt{18})} = 0.354$$

11) T' (0.354) is less than T (0.635), therefore the median baseline pollution loading is not exceeded.

2.2 Procedure B

2.2.1 Calculation of Single Observation Limit

- 1) The number of baseline observations, $n = 18$.
- 2) The natural log-transformed baseline observations are place in order from smallest to largest.
[-5.30, -3.51, -3.22, -3.10, -2.76, -2.75, -2.24, -1.83, -1.25, -0.90, -0.59, -0.50, -0.40, -0.11, 0.37, 0.59, 0.65, 1.02]
- 3) The mean of the 18 log-transformed observations, E_y , is -1.44.
- 4) Given the number of observations, the following equation is used to find A.

$$A = \frac{1}{1 - (\frac{2}{18} * 0.5)} = 1.06$$

- 5) The factor A is then used to calculate S_y^2 .

$$S_y^2 = 1.06 * \sum \frac{[(y_i - (-1.44))^2]}{17} = 3.24$$

- 6) To find E_x , the following equation is used:

$$E_x = \exp [(E_y) + (0.5 (S_y^2))]$$

$$E_x = \exp [(-1.44) + (0.5 * 3.24)] = \exp [0.18] = 1.20$$

- 7) Using the values that have been calculated the Single Observation Limit is found.

$$L_{so} = \exp [(-1.44) + (2.3263 * \sqrt{3.24})] = \exp [2.75] = 15.60$$

- 8) All of the monitoring observations are below 15.60, therefore the Single Observation Limit was not exceeded.

2.2.2 Warning Level

1) The Single Observation Warning Limit (WL_{so}) is determined using the following equation:

$$WL_{so} = \exp [(-1.44) + (1.6449 * \sqrt{3.24})] = \exp[1.52] = 4.58$$

2) All of the monitoring observations are below 4.58, therefore the Single Observation Warning Level is not exceeded.

2.2.3 Calculation of Cusum limit

1) The number of monitoring observations, $n = 18$.

2) The log-transformed monitoring observations are listed and labeled, in order of collection.

Natural Log-Transformed Monitoring Observations		Natural Log-Transformed Monitoring Observations	
Y_1	-0.63	$Y_{(10)}$	-0.19
Y_2	-3.22	$Y_{(11)}$	1.12
Y_3	0.04	$Y_{(12)}$	-0.54
Y_4	-3.41	$Y_{(13)}$	0.17
Y_5	-3.51	$Y_{(14)}$	-0.67
Y_6	-1.47	$Y_{(15)}$	-3.08
Y_7	-0.34	$Y_{(16)}$	-0.37
Y_8	-1.43	$Y_{(17)}$	-0.46
Y_9	-0.94	$Y_{(18)}$	-0.99

3) Using the values for E_y and S_y^2 from the L_{so} calculations, the value for K can be determined using the following equation:

$$K = [(E_y) + 0.25(\sqrt{S_y^2})]$$

$$K = (-1.44) + 0.25 * 1.8 = -0.99$$

- 4) The values for C_t can be determined using the following equation:

$$C_t = C_{t-1} + (Y_n - K) \text{ for example}$$

$$C_1 = -0.63 - (-0.99) = 0.36$$

The values for C_t are given in the table for each collection time t . Negative C_t values are replaced with 0, as shown in parentheses.

t	C_t	t	C_t
1	0.36	10	1.06
2	-1.87 (0)	11	3.17
3	1.03	12	3.62
4	-1.39 (0)	13	4.78
5	-2.52 (0)	14	5.10
6	-0.48 (0)	15	3.01
7	0.65	16	3.63
8	0.21	17	4.16
9	0.26	18	4.16

- 5) The baseline pollution Cusum Single Observation Limit, H , can be determined using the following equation:

$$H = 8.0 * (\sqrt{S_y^2})$$

$$H = 8.0 * (1.8) = 14.4$$

- 6) All values for C_t are below 14.4, therefore the baseline pollution Cusum Single Observation Limit was not exceeded.

2.2.4 Cusum Warning Level

- 1) The number of monitoring observations, $n = 18$.

- 2) The following equation can be used to determine K_w :

$$K_w = [(E_y) + 0.5(\sqrt{S_y^2})]$$

$$K_w = [(-1.44) + 0.5(1.8)] = -0.54$$

- 3) The values for W_t can be determined using the following equation:

$$W_t = W_{t-1} + (Y_n - K_w) \text{ for example}$$

$$W_1 = 0 + (-0.06 - (-0.0355)) = -0.0245 (0)$$

The limit values, W_t , are given in the table for each collection time t . Negative W_t values were replaced with 0, as shown in parentheses, and in the equation above.

t	W_t	t	W_t
1	-0.09 (0)	10	0.35
2	-2.68 (0)	11	2.01
3	0.58	12	2.01
4	-2.29 (0)	13	2.72
5	-2.97 (0)	14	2.59
6	-0.93 (0)	15	0.05
7	0.20	16	0.22
8	-0.69 (0)	17	0.30
9	-0.40 (0)	18	-0.15 (0)

- 4) The baseline pollution Cusum Warning Level, H_w , can be determined using the following equation:

$$H_w = 3.5 * (\sqrt{S_y^2})$$

$$H_w = 3.5 * (1.8) = 6.3$$

- 5) All values for W_t are below 6.3, therefore the baseline pollution Cusum Warning Level was not exceeded.

2.3 Annual Comparisons

2.3.1 Wilcoxon-Mann-Whitney test

Instructions for the Wilcoxon-Mann-Whitney test are given in Conover (1980), cited in Figure 3.2b.

- 1) When using both baseline and monitoring data, $n = 18$ and $m = 18$.

- 2) The baseline and monitoring observations are listed in order of collection, and ranked as follows:

Baseline Observations		Monitoring Observations	
0.030	2.5	0.530	20
0.005	1	0.040	6
1.915	34	1.040	30
0.673	25	0.033	4
0.064	10	0.030	2.5
0.063	9	0.230	13
0.607	23	0.710	27
0.553	21	0.240	14
0.286	15	0.390	17
0.106	11	0.830	28
0.406	18	3.050	36
1.447	32	0.580	22
0.900	29	1.180	31
0.040	5	0.510	19
2.770	35	0.046	8
1.803	33	0.690	26
0.160	12	0.630	24
0.045	7	0.370	16

The value of 0.030 was obtained for more than one observation. The ranking displayed is the average of 2 and 3 (2.5).

- 3) The sum of the 18 baseline ranks, $S_n = 322.5$.
- 4) From the table in section 1.3.1 of this appendix, the critical value (C) for 18 baseline and 18 monitoring observations is 281.
- 5) S_n (322.5) is greater than the critical value for C (281). Therefore, according to the Wilcoxon-Mann-Whitney test, the monitoring observations did not exceed the baseline pollution loading.

